Lecture 8
IR, Spring '24
Andrea Bajcsy

Last time:

- ☐ human behavior prediction (BP)
  - ☐ physics-based
  - ☐ planning-based

This time:

- ☐ intent inference
  - ☐ MLE vs. Bayesian
- ☐ data-driven BP

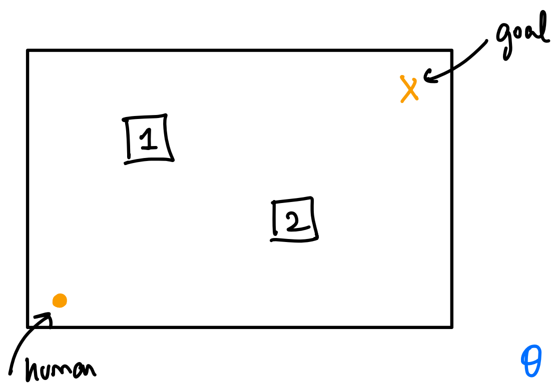
Last time, we ended by deriving a Maximum Entropy model of how likely diff. human traj. (i.e. demonstrations) are, given a reward parameter θ:

$$\xi := \{(s_1, a_1), (s_2, a_2) \dots (s_T, a_T)\}$$

$$P(\xi \mid \theta) = \frac{e^{\boxed{\theta^T \phi(\xi)}}}{\int e^{\theta^T \phi(\bar{\xi})} d\bar{\xi}}$$

$\equiv R_H(\xi; \theta)$ linear combo of weights $\theta \in \mathbb{R}^k$ and features $\phi : \Xi \to \mathbb{R}^k$. Assume ϕ is given / known but θ is UNKNOWN!

ex.



$$R_H(\xi; \theta) = \theta^T \phi(\xi)$$

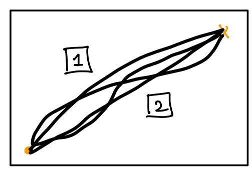
$$\phi(\xi) = \begin{bmatrix} \sum_{t=1}^{T} \text{distToObs}_1(x_t) \\ \sum_{t=1}^{T} \text{distToObs}_2(x_t) \\ \sum_{t=1}^{T} \text{distToGoal}(x_t) \end{bmatrix}$$

θ just combines these!

If we want to simulate human behavior under a particular reward param. we just sample from the distribution:

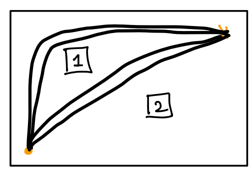
human cares about being CLOSE to 1, 2, & get to goal

$$\xi \sim P(\xi \mid \theta = [-1, -1, -10])$$



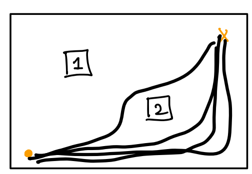
H only cares about 1

$$\xi \sim P(\xi \mid \theta = [-1, 0, -10])$$



far from 1 but close 2

$$\xi \sim P(\xi \mid \theta = [1, -1, -10])$$



Another important (but often overlooked) parameter of this model (Boltzmann model) is the temperature parameter β. → rationality param.

$$P(\xi \mid \theta) = \frac{e^{\beta \theta^T \phi(\xi)}}{\int e^{\beta \theta^T \phi(\bar{\xi})} d\bar{\xi}}$$

Here, $\beta \in [0, \infty)$ controls how optimally the human is expected to behave under the reward function $R_H(\cdot)$.

β

$0 \longrightarrow \infty$

$\beta = 0$

$\Rightarrow P(\xi \mid \theta) = \text{Unif}(\xi)$

$\beta \to \infty$

$\Rightarrow P(\xi \mid \theta) = \text{Dirac}(\xi^*)$

$\xi^* = \underset{\xi}{\arg\max}\ R_H(\xi; \theta)$

Ok, now we understand how the reward param θ of our predictive model influences our expectations of the human's behavior (as well as β).

But, how do we <u>infer</u> the θ from data?

TWO APPROACHES:-

1) $\boxed{\text{MLE}}$ Given $\xi_D \in D$ dataset of human trajectories.

$$\hat{\theta} = \underset{\theta}{\arg\max} \sum_{\xi_D \in D} \log P(\xi_D \mid \theta)$$

$$= \underset{\theta}{\arg\max} \sum_{\xi_D \in D} \log\left(e^{\theta^T \phi(\xi_D)}\right) - \log\left(\int e^{\theta^T \phi(\xi)} d\xi\right)$$

$$= \underset{\theta}{\arg\max} \underbrace{\sum_{\xi_D \in D} \theta^T \phi(\xi_D) - \log\left(\int e^{\theta^T \phi(\xi)} d\xi\right)}_{:= \mathcal{L}(\theta)}$$

$$\nabla_\theta \mathcal{L}(\theta) = \sum_{\xi_D \in D} \phi(\xi_D) - \frac{\partial}{\partial \theta} \log\left(\int e^{\theta^T \phi(\xi)} d\xi\right)$$

$$= \sum_{\xi_D \in D} \phi(\xi_D) - \frac{1}{\int e^{\theta^T \phi(\xi)} d\xi} \int \phi(\bar{\xi})\, e^{\theta^T \phi(\bar{\xi})} d\bar{\xi}$$

$$= \sum_{\xi_D \in D} \phi(\xi_D) - \int \phi(\bar{\xi}) \underbrace{\frac{e^{\theta^T \phi(\bar{\xi})}}{\int e^{\theta^T \phi(\xi)} d\xi}}_{:= P(\xi \mid \theta)} d\bar{\xi}$$

$$= \sum_{\xi_D \in D} \phi(\xi_D) - \int \phi(\bar{\xi})\, P(\bar{\xi} \mid \theta)\, d\bar{\xi}$$

$$= \sum_{\xi_D \in D} \phi(\xi_D) - \mathbb{E}_{\xi \sim P(\xi \mid \theta)}\left[\phi(\xi)\right]$$

empirical feature count of data ξ_D

expected feature count induced by current θ param.

This is an intuitive gradient update rule to minimize difference between the demonstrated feature counts & the expected feature counts under the estimated traj. distribution.

$$\hat{\theta}_{i+1} = \hat{\theta}_i + \alpha \overbrace{\left(\sum_{\xi_D \in D} \phi(\xi_D) - \mathbb{E}_{\xi \sim P(\xi \mid \theta)}\left[\phi(\xi)\right]\right)}^{\nabla_\theta \mathcal{L}(\theta)}$$

2) $\boxed{\text{Bayesian Inference}}$ Compute a posterior distribution over θ.

likelihood func :) $\propto e^{\theta^T \phi(\xi)}$

$$P(\theta \mid \xi_D) = \frac{\boxed{P(\xi_D \mid \theta)}\, P(\theta)}{\int P(\xi_D \mid \bar{\theta})\, P(\bar{\theta})\, d\bar{\theta}} \quad \text{// Bayes' Rule}$$

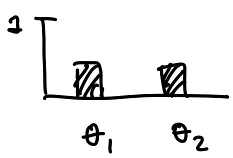
prior

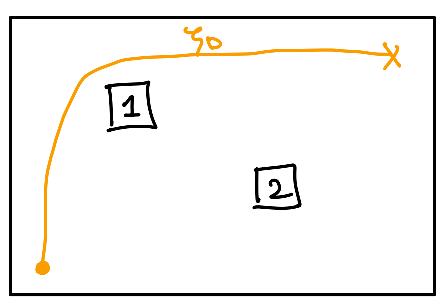
Assume: $\theta \in \Theta = \left\{ \underset{\theta_1}{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}, \underset{\theta_2}{\begin{bmatrix} -1 \\ 1 \end{bmatrix}} \right\}$

stay far from $\boxed{1}$ stay close to $\boxed{2}$

stay close $\boxed{1}$ far from $\boxed{2}$

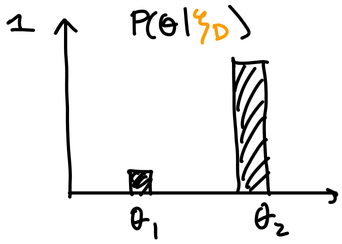
$P(\theta) = \text{Unif}(\Theta)$





$$P(\theta|\xi_D) = \frac{P(\xi_D|\theta)\,P(\theta)}{\int P(\xi_D|\bar{\theta})\,P(\bar{\theta})\,d\bar{\theta}}$$

$\propto e^{\theta^T\phi(\xi_D)}$



$\boxed{\theta = \theta_1}$ ⇒ it wants to stay far from $\boxed{1}$ and close to $\boxed{2}$
⇒ ξ_D UNLIKELY UNDER θ_1

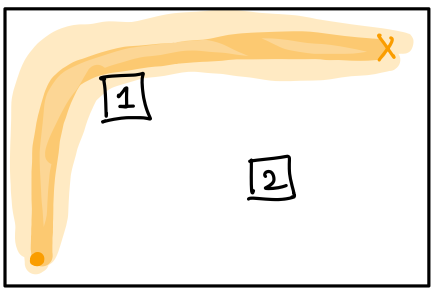
$\boxed{\theta = \theta_2}$ ⇒ it wants to stay close to $\boxed{1}$ and far from $\boxed{2}$
⇒ ξ_D MORE LIKELY under θ_2

$$P(\xi_D|\theta=\theta_1) < P(\xi_D|\theta=\theta_2)$$

This uncertainty / estimate θ can be used @ prediction time:

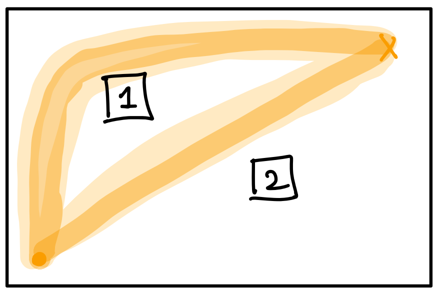
$\boxed{\text{MLE}}$ inferred $\hat{\theta}$

$$P(\xi|\hat{\theta}) \propto e^{\hat{\theta}^T\phi(\xi)}$$



$\boxed{\text{Bayesian}}$ have my belief $b(\theta) := P(\theta|\xi_D)$

$$P(\xi) = \sum_{\theta\in\Theta} b(\theta)\cdot P(\xi|\theta) \qquad \propto e^{\theta^T\phi(\xi)}$$



if there was uncertainty in the dataset over true θ, then your predictions "hedge" against this.

DATA / PATTERN-DRIVEN BEHAVIOR PREDICTION

Let's parameterize our predictive model via θ but θ isn't reward param, it's now the weights of a NN:

$$\mathcal{P}_\theta(x^{0:t}) \longmapsto \hat{x}^{t+1:T}$$

prediction of future states

simplest first pass.

INPUT OUTPUT

$(x^{0:t}, x^{t+1:T})$

$$\mathcal{D} = \{\xi_1, \xi_2, \ldots \xi_N\}$$

Solve:

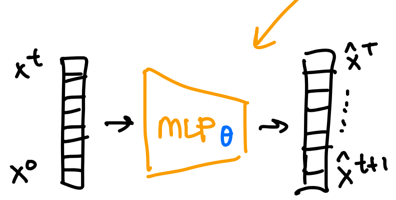
mean squared error

INPUT

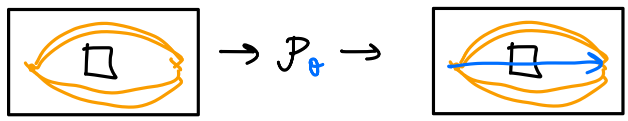
$$\min_\theta \sum_{\xi\in\mathcal{D}} \text{MSE}\big(\underbrace{\mathcal{P}_\theta(x^{0:t})}_{\rightarrow\, \hat{x}^{t+1:T}},\ x^{t+1:T}\big)$$

Deterministic Regression Problem!

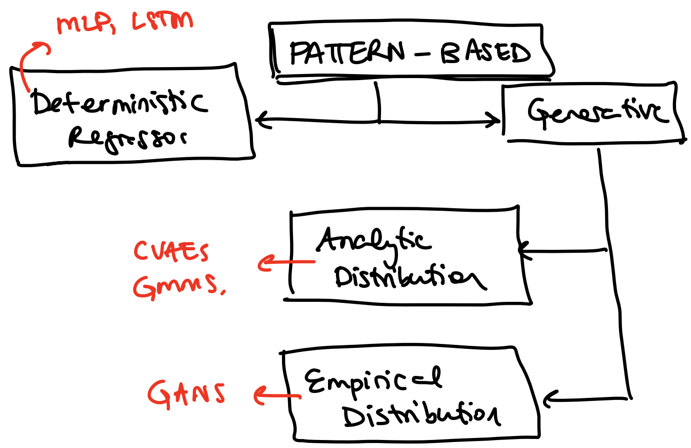
e.g.



Once again, we have trouble w/ multimodality here!

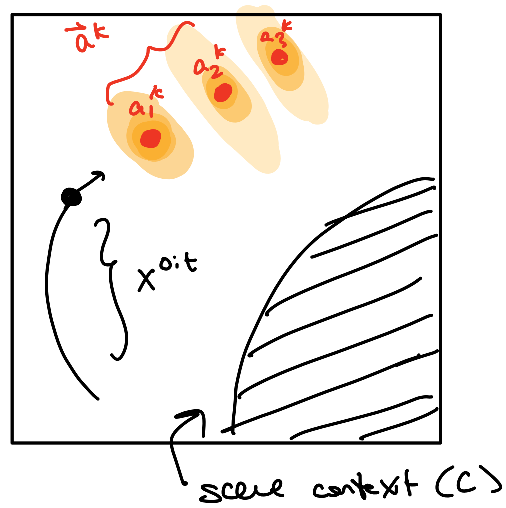


ENTER: generative models!
(e.g. GANs, CVAEs, GMMs...)



A common structure in SOTA human BP looks something like this:

- let $\vec{x}$ as history of states AND the scene context up to $[0 \dots t]$
- let $\vec{s} = [s_{t+1}, \dots s_T]$ be prediction



Multipath, WAYMO

$\vec{a}^k = [a_1^k, a_2^k \dots]$

\# of intent traj.

1) Discrete set of INTENTS: $\mathcal{A} = \{\vec{a}^k\}_{k=1}^{K}$

2) Uncertainty over intent:

$$P(\vec{a}^k \mid \vec{x}) = \frac{e^{\phi_k(\vec{x})}}{\sum_i e^{\phi_i(\vec{x})}}$$

$\phi_k : \mathbb{R}^{d(\vec{x})} \to \mathbb{R}$
output of NN

3) Uncertainty over state given intent:

$$P(s_t^k \mid \vec{a}^k, \vec{x}) = \mathcal{N}(s_t^k \mid a_t^k + \mu_t^k(\vec{x}), \Sigma_t^k(\vec{x}))$$

predicted by model!

4) prediction:

$$P(\vec{s} \mid \vec{x}) = \sum_{k=1}^{K} P(\vec{a}^k \mid \vec{x}) \prod_{t=1}^{T} P(s_t^k \mid \vec{a}^k, \vec{x}) \Big\} \text{ GMM model}$$

mixture weights.

⇒ trained via max. log likelihood of driving/behavior trajs.